Information-rich sampling of time-varying signals
PhD Thesis

Neeraj Kumar Sharma
Advisor: Prof. T. V. Sreenivas
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore-560012

Abstract

Confrontation with signal non-stationarity is a rule rather than an exception in the analysis of natural signals, such as speech, animal vocalization, music, bio-medical, atmospheric, ans seismic signals. Interestingly, our auditory system analyzes signal non-stationarity to trigger our perception. It does this with a performance which is unparalleled when compared to any man-made sound analyzer. Non-stationary signal analysis is a fairly challenging problem in the expanse of signal processing. Conventional approaches to analyze non-stationary signals are based on short-time quasi-stationary assumptions. Typically, short-time signal segments are analyzed using one of several transforms, such as Fourier, chirplets, and wavelets, with a predefined basis. However, the quasi-stationary assumption is known to be a serious limitation in recognizing fine temporal and spectral variations in natural signals. An accurate analysis of embedded variations can provide for more insightful understanding of natural signals.

Motivated from the sensory mechanisms associated with the peripheral auditory system, this thesis proposes an alternate approach to analyze non-stationary signals. The approach builds on the intuition (and findings from auditory neuroscience literature) that a sequence of zero-crossings (ZCs) of a sine-wave provides its frequency information. Building over this, we hypothesize that sampling an arbitrary signal at some signal specific time instants, instead of uniform Nyquist-rate sampling, can obtain a compact and informative dataset for representation of the signal. The information-richness of the dataset can be quantified by the accuracy to characterize the time-varying attributes of the signal using the sample dataset.

We systematically analyze this hypothesis for synthetic signals modeled by time-varying sinusoids and their additive mixtures. A restricted but rich class of non-stationary signals can be modeled using time-varying sinusoids. These sinusoids are characterized by their instantaneous-amplitude (IA) and instantaneous-frequency (IF) variations. It is shown that using ZCs of the signal and its higher-order derivatives, referred to as higher-order ZCs (HoZCs), we can obtain an accurate estimate of IA and IF variations of the sinusoids contained in the signal. The estimation is verified on synthetic signals and natural signal recordings of vocals and birdsong. On comparison of the approach with empirical mode decomposition, a popular technique for non-stationary signal analysis, and we show that the proposed approach has both improved precision and resolution.

Building on the above finding on information-richness in the HoZCs instant, we evaluate signal reconstruction using this dataset. The sampling density of this dataset is time-varying in a manner adapting to the temporally evolving spectral content of the signal. Reconstruction is evaluated for speech and audio signals. It is found that for the same number of captured samples, HoZCs corresponding to the first derivative of the signal (extrema samples) provide maximum information compared to other derivatives. This is found to be true even in a comparison of signals reconstructed from an equal number of randomly sampled measurements.

Based on these ideas we develop an analysis-modification-synthesis technique for a purely non-stationary modeling of speech signals. This is unlike the existing quasi-stationary analysis techniques. Instead, we propose to model the time-varying quasi-harmonic nature of speech signals. The proposed technique is not constrained by signal duration which helps to avoid blocking artifacts, and at the same time also provides fine temporal resolution of the time-varying attributes. The objective and subjective evaluations show that the technique has better naturalness post modification. It allows controlled modification of speech signals, and can provide for synthesizing novel speech stimuli for probing perception.